ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 11, November 2019

Multi-Keyword Search System on Cloud Computing

Nikhil Shirpurkar[1], Vipul Gore[1], Aniket Jagtap[1], Viraj N Jamdar[1], Prof. Amol Lachake[2]

U.G. Student, Department of Computer Engineering, Dr. DYPSOET, Pune, Maharashtra, India[1]

Associate Professor, Department of Computer Engineering, Dr. DYPSOET, Pune, Maharashtra, India [2]

ABSTRACT: The expanding prominence of distributed computing, an ever increasing number of information proprietors are persuaded to outsource their information to cloud servers for extraordinary accommodation and lessened cost in information administration. In any case, delicate information ought to be encoded some time recently outsourcing for security necessities, which obsoletes information usage like catchphrase based archive recovery. In this paper, we introduce a protected multi-watchword positioned seek conspire over scrambled cloud information, which at the same time underpins dynamic refresh operations like erasure and addition of records. In particular, the vector space show and the generally utilized TF IDF demonstrate are joined in the record development and inquiry age. We develop an extraordinary tree-based file structure and propose an "Eager Depth-first Search" calculation to give productive multi-catchphrase positioned look. The safe KNN calculation is used to encode the record and inquiry vectors, what's more, in the interim guarantee exact pertinence score estimation between scrambled list and inquiry vectors. So as to oppose measurable assaults, apparition terms are added to the record vector for blinding query items. Because of the utilization of our extraordinary tree-based record structure, the proposed plan can accomplish sub-direct pursuit time and manage the cancellation and inclusion of archives adaptable. Broad tests are led to show the effectiveness of the proposed conspires.

I. INTRODUCTION

Distributed computing has been considered as another model of big business IT foundation, which can sort out gigantic asset of registering, stockpiling and applications, and empower clients to appreciate omnipresent, advantageous and on demand organize access to a common pool of configurable figuring assets with extraordinary proficiency and negligible monetary overhead [1]. Pulled in by these engaging highlights, the two people and endeavors are spurred to outsource their information to the cloud, rather than obtaining programming and equipment to deal with the information them. Regardless of the different focal points of cloud administrations, outsourcing delicate data, (for example, messages, individual wellbeing records, organization back information, government reports, and so on.) to remote servers brings security concerns. The cloud specialist co-ops (CSPs) that keep the information for clients may get to clients delicate data without approval. A general way to deal with secure the information privacy is to scramble the information before outsourcing [2]. Be that as it may, this will cause a tremendous cost as far as information ease of use. For instance, the current systems on catchphrase based data recovery, which are generally utilized on the plaintext information, can't be specifically connected on the scrambled information. Downloading every one of the information from the cloud and unscramble locally is clearly unfeasible.

Keeping in mind the end goal to address the above issue, specialists have outlined some broadly useful arrangements with completely homomorphism encryption [3] or neglectful RAMs [4]. Be that as it may, these techniques are not down to earth because of their high computational overhead for both the cloud disjoin and client. On the opposite, more useful exceptional reason arrangements, for example, accessible encryption (SE) plans have made particular commitments as far as effectiveness, usefulness and security. Accessible encryption plans empower the customer to store the scrambled information to the cloud and execute watchword look over ciphertext space. Up until this point, bounteous works have been proposed under various danger models to accomplish different look usefulness, for

DOI: 10.15680/IJIRCCE.2019. 0711012

example, single watchword seek, likeness look, multi-watchword Boolean pursuit, positioned seek, multi-watchword positioned look, and so forth. Among them, multi-catchphrase positioned look accomplishes increasingly consideration for its pragmatic appropriateness. As of late, some unique plans have been proposed to help embedding's and erasing operations on report gathering. These are noteworthy works as it is exceptionally conceivable that the information proprietors need to refresh their information on the cloud server. However, few of the dynamic plans bolster proficient multi-catchphrase positioned look.

II. Literature Survey

[1] K. Bilal, M. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi Keyword Ranked Search over Encrypted Cloud Data", April 2011

In this paper, we study a semantic multi-keyword ranked search scheme over the encrypted cloud data, which simultaneously meets a set of strict privacy requirements. Firstly, utilize the "Latent Semantic Analysis" to reveal relationship between terms and documents. The latent semantic analysis takes advantage of implicit higher-order structure in the association of terms with documents ("semantic structure") and adopts a reduced - dimension vector space to represent words and documents. Thus, the relationship between terms is automatically captured. Secondly, our scheme employ secure "k-nearest neighbor (k-NN)" to achieve secure search functionality.

[2] C. Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data", 2012

In this paper, authors define and solve the problem of secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by enabling search result relevance ranking instead of sending undierentiated results, and further ensures the file retrieval accuracy. Specically, explore the statistical measure approach, i.e., relevance score, from information retrieval to build a secure searchable index, and develop a one-to-many order-preserving mapping technique to properly protect those sensitive score information. The resulting design is able to facilitate efficient server-side ranking without losing keyword privacy.

[3] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data"

In this paper, for the first time define and solve the problem of effective yet secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by returning the matching files in a ranked order regarding to certain relevance criteria (e.g., keyword frequency), thus making one step closer towards practical deployment of privacy-preserving data hosting services in Cloud Computing. The first give a straightforward yet ideal construction of ranked keyword search under the state-of-the-art searchable symmetric encryption (SSE) security definition, and demonstrate it's in efficiency.

[4] C. Gentry and S. Halevi, "Implementing gentrys fully homomorphic encryption scheme", 2011

The key-generation method for the underlying somewhat homomorphic encryption, that does not require full polynomial inversion. This reduces the asymptotic complexity from (n2.5) to (n1.5) when working with dimension-n lattices (and practically reducing the time from many hours/days to a few seconds/minutes). Other optimizations include a batching technique for encryption, a careful analysis of the degree of the decryption polynomial, and some space/time trade-os for the fully-homomorphic scheme.

[5] C. O rencik and E. Savas, "Efficient and secure ranked multi-key word search on encrypted cloud data", Mar 2012

In this paper, a practical privacy-preserving ranked keyword search scheme based on PIR that allows multi-keyword queries with ranking capability. The proposed scheme increases the security of the keyword search scheme while still satisfying efficient computation and communication requirements. To the best of our knowledge the majority of previous works are not efficient for assumed scenario where documents are large files.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 11, November 2019

III. PROPOSED ALGORITHM

Propose the first expressive SE scheme in the public-key setting from bilinear pairings in prime order groups. As such, our scheme is not only capable of expressive multi-keyword search, but also significantly more efficient than existing schemes built in composite-order groups. Using a randomness splitting technique, our scheme achieves security against offline keyword dictionary guessing attacks to the cipher texts. Moreover, to preserve the privacy of keywords against offline keyword dictionary guessing attacks to trapdoors, we divide each keyword into keyword name and keyword value and assign a designated cloud server to conduct search operations in our construction. Formalize the security definition of expressive SE, and formally prove that our proposed expressive SE scheme is selectively secure in the standard model.

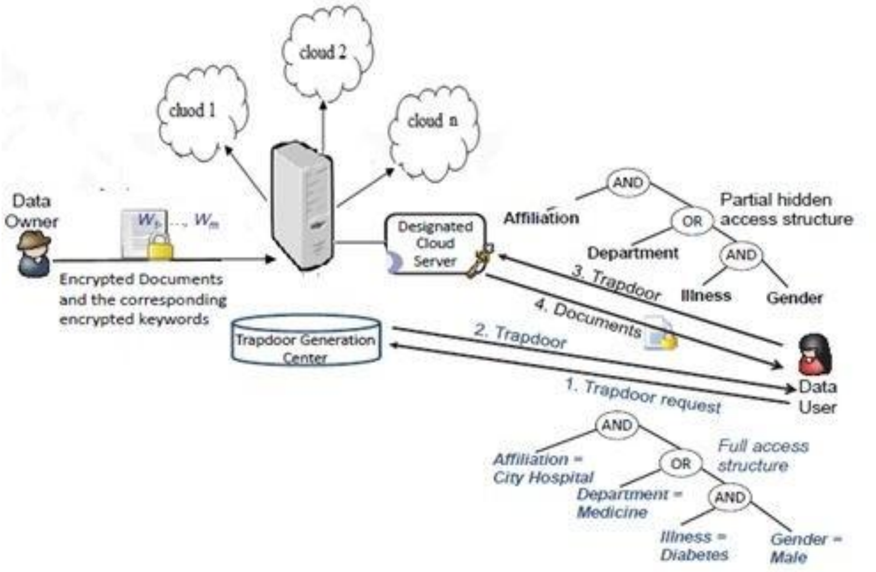

Fig 1: System Architecture

Implement our scheme using a rapidly prototyping tool called Charm, and conduct extensive experiments to evaluate its performance. Our results confirm that the proposed scheme is sufficiently efficient to be applied in practice. A trusted trapdoor generation center who publishes the system parameter and holds a master private key and is responsible for trapdoor generation for the system, data owners who outsource encrypted data to a public cloud, data users who are privileged to search and access encrypted data, and a designated cloud server who executes the keyword search operations for data users. Distributed computing is used to refer to any large collaboration in which many individual personal computer owners allow some of their computer's processing time to be put at the service of a large problem. In our system the each cloud admin consist of data blocks. The cloud users upload the data into multi cloud. Cloud computing environment is constructed based on open architectures and interfaces; it has the capability to incorporate multiple internal and/or external cloud services together to provide high interoperability. We call such a distributed cloud environment as a multi-Cloud. A multi-cloud allows clients to easily access his/her resources remotely through interfaces. Division of Data in the Cloud for Optimal Performance and Security that collectively approaches the security and performance issues. In the division methodology, we divide a file into fragments data over the cloud nodes. Each of the nodes stores only a single fragment of a particular data file that ensures that even in case of a successful attack, no meaningful information is revealed to the attacker. Our main aim is to build a system where Optimization of Performance and Security of data in the cloud is enhanced. We implemented division of data in the cloud for excellent performance and security; we correlated the particular data division of a file into fragments.

DOI: 10.15680/IJIRCCE.2019. 0711012

IV. CONCLUSION AND FUTURE WORK

In this paper, a safe, effective and dynamic pursuit conspire is proposed, which underpins not just the precise multi-keyword positioned look yet in addition the dynamic cancellation and inclusion of archives. We develop a unique watchword adjusted paired tree as the record, and propose a "Ravenous Profundity initially Search" calculation to acquire better proficiency than direct hunt. Also, the parallel pursuit process can be done to additionally lessen the time cost. The security of the plan is ensured against two risk models by utilizing the protected KNN calculation. Trial comes about illustrate the effectiveness of our proposed conspire. There are as yet many test issues in symmetric SE plans. In the proposed plot, the information proprietor is dependable for producing refreshing data and sending them to the cloud server. Along these lines, the information proprietor needs to store the decoded list tree and the data that are important to recalculate the IDF esteems. Such a dynamic information proprietor may not be extremely reasonable for the distributed computing model. It could be an important however troublesome future work to outline a dynamic accessible encryption plot whose refreshing operation can be finished by cloud server just, in the interim saving the capacity to help multi-watchword positioned look.

REFERENCES

1. C. Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data," IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 8, pp. 1467–1479, Aug. 2012.
2. B. Zhang and F. Zhang, "An efficient public key encryption with conjunctive-subset keywords search," J. Netw. Comput. Appl., vol. 34, no. 1, pp. 262–267, 2011.
3. B. Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preserving multi-keyword fuzzy search over encrypted data in the cloud," in Proc.IEEE INFOCOM, 2014, pp. 2112–2120.
4. C. Wang, K. Ren, S. Yu, and K. M. R. Urs, "Achieving usable and privacy-assured similarity search over outsourced cloud data," in Proc. IEEE INFOCOM, 2012, pp. 451–459.
5. M. Kuzu, M. S. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data," in Proc. IEEE 28th Int. Conf. Data Eng., 2012, pp. 1156–1167.
6. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," in Proc. IEEE INFOCOM, Apr. 2011, pp. 829–837.
7. W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H. Li, "Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking," in Proc. 8th ACM SIGSAC Symp. Inf., Comput. Commun. secur., 2013, pp. 71–82.
8. C. Orencik, M. Kantarcioglu, and E. Savas, "A practical and secure multi-keyword search method over encrypted cloud data," in Proc. IEEE 6th Int. Conf. Cloud Comput., 2013, pp. 390–397.
9. W. Zhang, S. Xiao, Y. Lin, T. Zhou, and S. Zhou, "Secure ranked multi-keyword search for multiple data owners in cloud computing," in Dependable Syst. Networks (DSN), IEEE 44th Annu. IEEE/IFIP Int. Conf., 2014, pp. 276–286.
10. S. Kamara, C. Papamanthou, and T. Roeder, "Dynamic searchable symmetric encryption," in Proc. ACM Conf. Comput. Commun. Secur., 2012, pp. 965–976.
11. S. Kamara and C. Papamanthou, "Parallel and dynamic searchable symmetric encryption," in Proc. Financ. Cryptography Data Secur., 2013, pp. 258–274.\
12. D. Cash, S. Jarecki, C. Jutla, H. Krawczyk, M.-C. Ro¸su, and M. Steiner, "Highly-scalable searchable symmetric encryption with support for boolean queries," in Proc. Adv. Cryptol, 2013, pp. 353–373.